

COMBINED CHARGES DIFFICULTY EVALUATION AND TESTEES LEVEL FOR INTELLECT APPRAISAL

¹S.Bhuvana, ²Lokesh Raj R.M, ²Sriram.S, ²Yuvaraj.P

¹Assistant Professor, ²Final year student, Department of CSE,
bhupreethi@gmail.com, lokeshrm8@gmail.com,

Received 20 March 2020 Received in revised form 08 April 2020 Accepted 10 April 2020
Available online 19 April 2020

ABSTRACT

In this paper, the combined charge difficulty evaluation and testees level for intellect appraisal are to find the employee performance. In testing the tests and task evaluation of different designation employee's performance outcome result is difficult to rank the outcome result. Whereby higher officials evaluate and provide feedback on employee job performance, including steps to improve or redirect activities as needed. Higher official documenting performance provides a basis for pay increases and promotions. In this project, higher official evaluates the lower official by assigning task and testees. The higher official will solve the difficulties of testing and ranking problem for appraisal. All these findings are useful to intelligence tests. Here the evaluation of each employee was calculated and view by the Bernoulli's Distribution algorithm

Keywords-Task evaluation, employee appraisal, ranking of employees

I. Introduction

TESTING the intelligence level of a human or an artificial intelligent system is an important yet hard problem. Since it is really hard to directly define intelligence, we usually indirectly characterize the intelligence level of the human or artificial intelligent systems according to their behaviors in some dedicated tests.

For example, we often test the intelligence levels of children based on their performance in some standard tasks designed by human experts. However, such methods may not work for some artificial intelligent systems, since we do not have a standard comparison level for tasks.

In this paper, we study the testing tasks evaluation and testees ranking problem. Tasks may have different difficulty levels, and testees may have different capabilities; however, we do not know in such a scope. Usually, we setup a series of tasks (e.g., traffic sign detection, vehicle detection, and pedestrian detection) for the vehicle and check whether it can successfully finish these tasks in time. However, we cannot accurately determine the difficulties of each

task. The vehicles that take part in the tests may also have noticeably various capabilities, so that we cannot judge the difficulty of each task simply by their performance.

Our objective is to simultaneously determine the relative difficulty level of each testing task and the relative capability of every testee, purely based on the test outcome of every testees on each task. Moreover, we assume that testees may have a probability to pass a certain task so as to allow certain uncertainty.

To solve this problem, we have designed two models to solve this problem. One model assumes that the test outcome follows a certain Bernoulli distribution, while the other assumes that the test outcome follows a certain Bernoulli distribution with the beta distribution-type *a priori* knowledge. We formulate the problem as likelihood estimation problems. And propose coordinate descent algorithm to solve the Parameters of the studied distributions. We show that the beta distribution-type *a priori* knowledge is useful, when we only carry out a limited number of tests due to time and financial budgets.

To better present our findings, the rest of this paper is arranged as follows. Sections II and III present the

problems and give two statistical learning models as well as the solving algorithm. Section IV provides an empirical numerical study. Section V discusses how to extend this model for more general cases. Finally, Section VI provides the conclusion.

This model is somewhat like the logistic principal component analysis model proposed. However, our model adds quite different assumptions on the latent structure of the statistical model.

Clearly, to derive an estimation of p_i and q_j is equivalent to finding a parameter set of p_i and q_j that maximize L . However, it is not a convex optimization problem for maximum likelihood estimation (MLE) problem

Summarizing the above analysis, we propose the following coordinate descent algorithm. The numerical tests show that this coordinate descent algorithm converges to the local optimal solution quickly. To increase the probability to hit the global optimal solution, we can run the following algorithm for several times with different initial values. Therefore, each iteration can increase or maintain the lower bound estimation of the log-likelihood. Consequently, the lower bound of the log-likelihood will finally converge to a local maximum

II. TESTING RESULTS

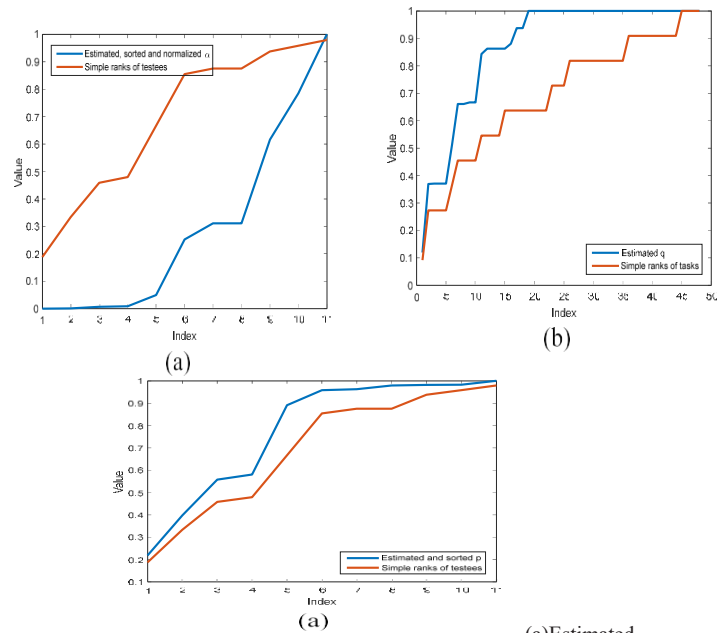
Here, we give a numerical example for intelligent vehicles testing. In November 2017, 11 teams took part in the vehicle detection competition supported by the National Natural Science Foundation of China. The competition consists of 48 vehicles to detect, and the testing results are shown in Fig. 1. Each column represents for a testee while each row represents for a task vehicle. Then, the element in the j th row and i th column stands for the binary detection result, where 1 means passed and 0 means failed. (worked out 46 tasks) is very close to team 7 (worked out 47 tasks), their capability differences should not be omitted. Intuitively, the most difficult task, i.e., task 7, and the most capable testee, i.e., team 7, located in the top right of each figure, should be distinguishable from other tasks and testees, especially nearby ones. Such distinguish ability implies the curves of estimated and sorted p_i and q_j should not be flat. Contrarily, the Bernoulli distribution with the beta distribution-type *a priori* knowledge model appropriately emphasizes

this point and, meanwhile, gives the prominence to the difficulty level of task 7, as shown in Fig. 3. Therefore, we can conclude that, when we have limited times of tests for a single task, the beta distribution-type *a priori* knowledge is needed

III. FURTHER DISCUSSION

Model Estimation Considering Missing Data

The afore mentioned statistical learning model (1) provides a basis for simultaneous tasks evaluation and testees ranking. However, in many situations,



and sorted p_i and the simple ranks of the testees

(b) Estimated and sorted q_j , and the simple ranks of the testees.

Do not have enough time and money to carry out tests on each task for every testee. To handle the parameter estimation with respect to missing observations, we sort to the classic expectation maximization (EM) algorithm. Suppose that the available part of $x_{i,j}$ can be written as x^{obs} and the missing part of $x_{i,j}$ can be written as x^{miss} . The detailed algorithm is given as follows. It is easy to prove that the above algorithm is a standard EM algorithm can converge to a local optimal solution

quickly. We need to run the above algorithm for several times with different initial values to increase the probability to hit the global optimal solution. The larger the missing ratio, the larger is the number of running times. Similarly, we can handle the cases when the beta distribution-type *a priori* knowledge is considered

Model Estimation Considering Multiple Experiments

The above models assume that every testee runs each test once. However, we can further relax this assumption and allow the i th testee to run the j th testing task for $y_{i,j}$ times and record the number of successfully finish. We can still derive an estimation of p_i and q_j by finding a parameter set of p_i and q_j to maximize the likelihood function

IV. CONCLUSION

In this paper, we discuss how to simultaneously determine the difficulty levels of testing tasks and the capabilities of testees, especially when heterogeneous tasks and testees are considered. We propose several statistical learning models to handle such problems and allow testees to have some certain uncertainties in finishing tasks. The numerical test results verify the effectiveness of the proposed learning models and also the solving algorithms. All these findings are useful to designs of tests for artificial intelligent systems.

It should be pointed out that the aforementioned algorithm only provides an up-to-time estimation of the tasks and testees, because the capability of a testee may grow with time and the relative difficulty of a task may also vary accordingly. Similar to the Elaborating system for chess and Go games [6], we need to continuously design new and tougher tasks to keep a proper understanding of the testees.

V. REFERENCES

- [1] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. London, U.K.: Pearson Education, 2010.
- [2] L. Li *et al.*, "Artificial intelligence test: A case study of intelligent vehicles," *Artif. Intell. Rev.*, vol. 50, no. 3, pp. 441–465, Oct. 2018.
- [3] R. J. Sternberg, *Beyond IQ: A Triarchic Theory of Human Intelligence*. Cambridge, U.K.: Cambridge Univ. Press, 1985.
- [4] N. Mackintosh, *IQ and Human Intelligence*, 2nd ed. Oxford, U.K.: Oxford Univ. Press, 2011.
- [5] S. Ohlsson, R. H. Sloan, G. Turán, and A. Urasky, "Measuring an artificial intelligence system's performance on a verbal IQ test for young children," *J. Experim., Theor. Artif. Intell.*, vol. 29, no. 4, pp. 679–693, 2017.
- [6] D. Silver *et al.*, "Mastering the game of Go without human knowledge," *Nature*, vol. 550, pp. 354–359, Oct. 2017.
- [7] L. Li, D. Wen, N.-N. Zheng, and L.-C. Shen, "Cognitive cars: A new frontier for ADAS research," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 1, pp. 395–407, Mar. 2012.
- [8] W. Huang, D. Wen, J. Geng, and N.-N. Zheng, "Task-specific performance evaluation of UGVs: Case studies at the IVFC," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 5, pp. 1969–1979, Oct. 2014.
- [9] L. Li, W.-L. Huang, Y. Liu, N.-N. Zheng, and F.-Y. Wang, "Intelligence testing for autonomous vehicles: A new approach," *IEEE Trans. Intell. Veh.*, vol. 1, no. 2, pp. 158–166, Jun. 2016.
- [10] A. I. Schein, L. K. Saul, and L. H. Ungar, "A generalized linear model for principal component analysis of binary data," in *Proc. 9th Int. Workshop Artif. Intell. Statist.*, Jan. 2003, p. 10.
- [11] L. Li, *Selected Applications of Convex Optimization*. Beijing, China: Springer, 2015.
- [12] A. K. Gupta and S. Nadarajah, Eds., *Handbook of Beta Distribution and Its Applications*. Boca Raton, FL, USA: CRC Press, 2004.
- [13] A. Dempster, N. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc. Ser. B*, vol. 39, no. 1, pp. 1–38, 1977.
- [14] M. E. Tipping and C. M. Bishop, "Probabilistic principal